

EXPRESS EV 36539686745PF 020124
10/528729

a/ppts

JCO6 Rec'd PCT/PTO 22 MAR 2005

METHOD AND DEVICE FOR MEASURING SIMILARITY BETWEEN
IMAGES

5 The invention relates to a method and device for measuring similarity between images.

10 The field of application is that of the calculation of the similarity between images. This similarity may be used within the framework of a search for images by similarity. It may also be used to estimate clusters of images that are close according to the similarity criterion used. It pertains for example to the construction of video summaries. This consists of a search for similar images in a base consisting of
15 images of a video sequence. The aim is to reduce the number of these images by retaining only a representative for a cluster of images that are decreed to be similar. It also pertains to indexation consisting in selecting index images on the basis of
20 their similarity to other images or quite simply in the search for specific images in the video sequence, on the basis of enquiry images.

25 Numerous studies are interested in the measurement of similarity between images on the basis of attributes assigned to the images. A certain number of them are now defined in the MPEG-7 standard. The invention is interested in particular in the so-called structural attributes. For example, in the latest version of the
30 standard, an attribute of contour type is defined. It is extracted locally from blocks of images on the basis of Gabor filters.

35 The attributes known from the prior art or proposed in the MPEG-7 standard and the algorithms for processing these attributes for measuring similarity, do not however make it possible, for certain types of sequences, to detect similar images in a satisfactory manner, in so far as they do not take account either of

the image in its entirety, or of the details in the image. With a view to clustering images extracted from a video, the similarity measurement must be able to convey the fact that several images correspond to one
5 and the same scene, characterized by a unity of place. An information item characterizing solely the details of the image is not sufficient to fulfil this task. On the other hand an information item characterizing the image solely in its entirety may prove to be unsuitable
10 in the case of partial modification of the staging of the scene, for example by modification of the background of the scene.

The aim of the invention is to alleviate the aforesaid
15 drawbacks. Its subject is a method of measuring similarity between images, characterized in that it performs, for each image, the following steps:

- segmentation of the image,
- classification of the segments as a function of
20 their orientation to give classes,
- calculation of a histogram of the number of segments as a function of class,
- calculation of a histogram of the number of pixels belonging to the segments of one and the same
25 class as a function of class,
- comparison of the histograms of each image to give a measurement of similarity.

According to a particular implementation, the process
30 is characterized in that it also calculates a histogram corresponding to the distribution of the segments about the centre of gravity of each class.

According to a particular embodiment, to calculate this
35 histogram, it performs a calculation of the standard deviation of the distances from the middles of the segments of a class to the centre of gravity of the class considered.

According to a particular implementation, the comparison of the histograms consists of a subtraction of the ordinates, class by class and of a sum, over the set of classes, of the values obtained for each class.

5

According to a particular implementation, the histograms are coded according to the MPEG-7 standard.

10 The invention also relates to a method of clustering images of a database, characterized in that the clustering is performed as a function of the measurements of similarity according to the method described above to give clusters of images.

15 The invention also relates to a method of creating video summaries, characterized in that it selects at least one of the images of at least one cluster calculated according to the above method.

20 The invention also relates to a method of video indexation, characterized in that it selects at least one of the images of at least one cluster calculated according to the above method, as indexation image.

25 The invention also relates to a device for measuring similarity between images, characterized in that it comprises a circuit for processing and for calculation of histograms receiving digital data defining these images so as to perform, for each of them, the following operations:

30

- segmentation of the image,
- classification of the segments as a function of their orientation to give classes,
- calculation of a histogram of the number of segments as a function of class,
- 35 - calculation of a histogram of the number of pixels belonging to the segments of one and the same class as a function of class,
- comparison of the histograms of each image to

give a measurement of similarity.

The main advantage of the invention is the implementation of powerful algorithms for
5 characterizing an image and measuring the similarity between images on the basis of these algorithms, by virtue of the joint use of attributes based on orientation of contours making it possible to
10 characterize the number of segments, their size and also their distribution as a function of orientation.

The global description of the image is obtained on the basis of the histograms of the orientations of the segments in the image, hence without taking account of
15 the position of these segments in the image, and the local description is obtained on the basis of a measure of centres of gravity of classes of segments, which corresponds to the relative positions of the various segments of one and the same class.

20 The use of structural attributes such as these makes it possible to characterize the image both locally and globally, thus improving the effectiveness in the search for similarity of images. Specifically, the
25 dimensions of the similar objects may be very different from one image to another, from one shot of the sequence to another.

Other features and advantages of the invention will
30 become clearly apparent in the following description given by way of nonlimiting example, and offered in conjunction with the appended figures which represent:

- Figure 1, a flowchart of the method according to the invention,
- 35 - Figure 2, a method of polygonal approximation,
- Figure 3, an exemplary search for similarity.

The method carries out a segmentation of the image. The "object" approach does not however exist here. It

involves determining segments in the image which are approximations of the actual contours in the image, for example on the basis of measurements of gradients, without worrying about the fact that they do or do not
5 belong to a particular object in the image.

Figure 1 represents a flowchart defining the main steps of the method of calculating histograms of attributes characterizing an image.

10

A first phase which is a preprocessing of the image consists of a detection of contours in the image so as to obtain a map of contours and a segmentation of these contours so as to obtain a map of segments. The second
15 phase carries out a calculation of the attributes for the detection of similarity.

An image to be processed is transmitted to a first step referenced 1 in the flowchart of Figure 1. This step of
20 the method carries out a calculation of gradients in the image. A gradients detection filter is applied to the image to give at each point a measurement of gradient. A map of vertical gradients and a map of horizontal gradients are thus calculated. The norm of
25 the gradients that is obtained on the basis of these maps, the square root of the sum of the squares of the horizontal and vertical gradient values attributed to a pixel, is also utilized.

30 Step 2 carries out a selection of the pixels on the basis of the gradient values calculated in the previous step, compared with thresholds. This selection is refined through that of the points of greatest contrast in a given neighbourhood window. The aim is to remove
35 thick contours by selecting, in the neighbourhood of a pixel, the pixels having the steepest gradients, by considering the values of the horizontal and vertical gradients. This also involves favouring the pixels neighbouring a pixel already selected as a contour. The

isolated contour points are eliminated. We obtain, at this stage, a binary map of contours, each pixel of the image is labelled as contour or otherwise.

- 5 A test of connectivity is performed on the binary map of contours so as to gather together the neighbouring pixels, so as to thus constitute contour lines. To do this, a procedure for labelling as adjoining components is implemented in the next step 3.

10

Two pixels are declared adjoining if there exists an uninterrupted path of contour pixels linking these two pixels. Contour pixels are thus "chained together" in such a way as to obtain a continuity line. The adjoining pixels are labelled, each pixel touching another pixel being marked with the same label. A list of adjoining components is obtained for each label, this list being the coordinates of the various pixels of a continuity line relating to a label.

20

To chain each contour together from one end to the other, it is necessary to eliminate multiple points, that is to say the junctions of contour lines. An algorithm tests for the presence of junctions and eliminates them. This step is carried out by tracking contours over each adjoining component identified. The traversal of the pixels belonging to a given adjoining component is done gradually. A combination of filters makes it possible to ascertain whether the current pixel corresponds to a junction between the main contour and a secondary contour. A pixel having been categorized as a junction is eliminated from the map of contour pixels, thus detaching the secondary contour from the main contour. The secondary contour is replaced in the list of adjoining components to be traversed. The tracking of contours is then continued along the main contour.

30

35

At this stage, we have a map of nonrectilinear

continuity lines.

The next step 4 of the preprocessing procedure is a polygonal approximation of these chained contours so as to approximate each contour line by a set of connected segments.

Figure 2 represents a method of polygonal approximation of a chained contour. For a given contour referenced 11, we search for the point e3 giving the maximum distance between the points of the contour and the segment formed by the ends of this contour e1 and e2. If this distance is greater than a threshold, the contour is approximated by two segments [e1, e3] and [e3, e2]. The procedure is repeated until the distances from the point to the segment are less than the fixed threshold.

On completion of this step, we have, for an image, a segment map consisting of a set of segments whose orientation can then be calculated. An example of a map of segments is represented at the top right of Figure 3.

The second phase consists of a calculation and a utilization of the attributes relating to this segment map.

Step 5 calculates, for each segment of the map of segments, the angle between its direction and the horizontal. The angle obtained, lying between 0 and 180 degrees, is then quantized to obtain a reduced number of categories or classes of angles. For example, the angles obtained are clustered into 36 classes, the quantization interval being 5 degrees or, stated otherwise, a class corresponds to a width of 5 degrees.

Histograms relating to three attributes are now calculated during step 6, utilizing the segment

orientation information obtained previously.

- a histogram of the number of segments as a function of orientation.

- 5 The various classes are represented as abscissa and the occurrences as ordinates.

10 - a histogram of the total number of pixels belonging to segments of like orientation as a function of their orientation.

The ordinates correspond to the sum of the lengths of the segments of the class, for a given class.

- 15 - a histogram corresponding to the distribution of the segments about the centre of gravity of each class, by calculating the standard deviation of the distances from the middles of the segments to the centre of gravity.

20 The centre of gravity G_i of class C_i is obtained by calculating the barycentre of the middles $M(s_j)$ of the segments s_j . The weight associated with each middle is equal to the number of pixels $n(s_j)$ making up the segment:

25

$$G_i = \frac{\sum_{s_j \in C_i} n(s_j) * M(s_j)}{\sum_{s_j \in C_i} n(s_j)}$$

The ordinates $H(i)$ correspond to the standard deviation of the Euclidean distances $d(G_i, M(s_j))$ between the barycentre or centre of gravity of the class C_i considered and the centre of each segment of the class:

30

$$H(i) = \sqrt{\frac{\text{card}(C_i) * \sum_{s_j \in C_i} d(G_i, M(s_j))^2 - (\sum_{s_j \in C_i} d(G_i, M(s_j)))^2}{\text{card}(C_i) * \text{card}(C_i)}}$$

where card is the cardinal function returning the number of segments of the class considered.

5 The distribution is put into the form of a histogram, that is to say of vectors of values, thus allowing its utilization within the framework of the MPEG7 standard which defines the coding of such histograms.

10 In order to more easily combine the information, the histograms are normalized, for example on the basis of the maximum values found.

15 The three histograms of the attributes characterizing the image are utilized for the similarity measurementments, step 7.

Three histograms are calculated for a first image which is for example an enquiry image and three other histograms for a second image which is an image in a database. The measurement of similarity between the two images I_1 and I_2 , that is to say the calculation of the distance between these images may then be performed by simple subtraction of the ordinates, class by class, for each type of histogram H^k of size m_k :

25

$$d_k(H_1^k, H_2^k) = \frac{\sum_i |H_1^k(i) - H_2^k(i)|}{m_k}$$

A similarity indicator $S(I_1, I_2)$ is for example the sum of the values obtained by these subtractions, for the whole set of histograms:

30

$$S(I_1, I_2) = \frac{\sum_{k=1,2,3} d_k(H_1^k, H_2^k)}{3}$$

The similarity measurement may be improved by comparing the class also with neighbouring classes of the histogram of the other image, with a weighting factor.

The histogram comparison uses for example a metric of cross quadratic type making it possible to circumvent the small variations in global orientation between two successive images, for example during a small rotation
5 of the camera.

An exemplary search for similar images is given in Figure 3, showing, at the top left, the image considered, its associated map of contours and of
10 segments, as well as the images recognized as the most similar in a base of 150 images arising from two different sequences.

The main characteristic of these histograms is
15 therefore of providing a description of the images which is both global and local. It is thus possible, by virtue of the global description, to differentiate globally structured images, for example images of towns characterized by orientations of horizontal and
20 vertical segments, from image is of countryside characterized by orientations of segments that are more random. However, it is also possible, by virtue of the local description, to differentiate locally structured images, part of the image is for example devoted to a
25 building, another to a wood, from less structured images, for example a house in the middle of the forest.

The calculation of the segments, in the previous
30 description, is performed over the values of luminance of the image. It is of course just as conceivable to perform such a calculation for each of the colour components of the image. If the calculations need more processing capacity, on the other hand the production
35 of histograms based on colour makes it possible to improve the effectiveness of the method.

Three histograms are utilized here to characterize the image. It is of course possible to utilize just one of

them or a combination of two of them to calculate the similarity of images, without departing from the field of the invention.

- 5 The segment orientations utilized may be the angles formed with reference straight lines other than the horizontal. The dimensions of the classes may be less than or greater than 5 degrees, depending on the calculation power or time available or the quality of
- 10 measurement desired.